

**ANALISIS *CLUSTER* KABUPATEN/KOTA DI RIAU
BERDASARKAN PENDAPATAN DOMESTIK REGIONAL
BRUTO (PDRB) TAHUN 2010**

TUGAS AKHIR

Diajukan sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Sains
pada Jurusan Matematika

oleh:

SERI DARIYANTI

10654004496



**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU
2011**

**ANALISIS *CLUSTER* KABUPATEN/KOTA DI RIAU
BERDASARKAN PENDAPATAN DOMESTIK
REGIONAL BRUTO (PDRB) TAHUN 2010**

**SERI DARIYANTI
NIM: 10654004496**

Tanggal Sidang : 28 Juni 2011
Periode Wisuda : Oktober 2011

Jurusan Matematika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Riau
Jl. HR. Soebrantas No.155 Pekanbaru

ABSTRAK

Penelitian ini membahas tentang analisis *cluster* dalam bidang pemerintahan yaitu mengelompokkan kabupaten/kota di Riau berdasarkan PDRB Tahun 2010. Dalam tugas akhir ini, data yang digunakan adalah data PDRB Tahun 2010 menurut lapangan usaha. Sebelum melakukan analisis *cluster*, data terlebih dahulu dianalisis menggunakan analisis komponen utama (AKU) dengan penguraian nilai singular. Ukuran ketakmiripan dalam analisis *cluster* untuk menentukan jarak menggunakan metode Euclid, sedangkan perbaikan jarak menggunakan metode pautan tunggal. Hasil analisis data menunjukkan bahwa terdapat 7 kelompok/*cluster*.

Kata kunci: *analisis cluster, analisis komponen utama, jarak euclid, metode pautan tunggal*

KATA PENGANTAR

Alhamdulillahirabbil'alamin, puji syukur senantiasa penulis ucapkan kehadirat Allah SWT, atas segala limpahan rahmat serta karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir ini dengan judul “ANALISIS *CLUSTER* KABUPATEN/KOTA DI RIAU BERDASARKAN PENDAPATAN DOMESTIK REGIONAL BRUTO (PDRB) TAHUN 2010”. Shalawat serta salam semoga senantiasa tercurahkan kepada junjungan nabi besar, Nabi Muhammad SAW, keluarga sahabat, serta umatnya.

Sudah menjadi sunatullah dalam hidup ini, setiap ada awal pasti ada akhir, tidak terkecuali dengan penulisan tugas akhir ini, yang merupakan akhir dari suatu proses berantai sejak dari awal duduk dibangku kuliah. Tugas akhir ini disusun sebagai syarat untuk memenuhi salah satu syarat dalam menyelesaikan studi Strata satu (S1) di UIN Suska Riau.

Penulis menyadari bahwa tidak sedikit yang terlibat dalam penulisan ini baik yang memberi bantuan moril maupun materil, oleh karena itu dengan kerendahan hati yang tulus penulis menyampaikan ucapan terima kasih yang setinggi-tingginya kepada kedua orang tua tercinta ayahanda (Saliono) dan ibunda (almh. Hamidah), yang membuat penulis mantap untuk terus melangkah. Selanjutnya ucapan terimakasih yang tak terima kasih kepada:

1. Bapak Prof. Dr. H. M. Nazir selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Ibu Dra. Hj. Yenita Morena, M.Si selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Ibu Yuslenita Muda, M.Sc selaku Ketua Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Ibu Ari Pani Desvina, M.Sc selaku Pembimbing yang telah banyak membantu, mengarahkan, mendukung dan membimbing penulis dalam penulisan tugas akhir ini.

5. Bapak Rado Yendra, M.Sc selaku Penguji I yang telah banyak membantu, memberikan kritikan dan saran serta dukungan dalam penulisan tugas akhir ini.
6. Ibu Fitri Aryani, M.Sc selaku Penguji II dan Koordinator Tugas Akhir yang telah banyak membantu, memberikan kritikan dan saran serta dukungan dalam penulisan tugas akhir ini.

Dalam penyusunan tugas akhir ini penulis telah berusaha semaksimal mungkin. Walaupun demikian tidak tertutup kemungkinan adanya kesalahan dan kekurangan baik dalam penulisan maupun dalam penyajian materi. Untuk itu penulis mengharapkan kritik dan saran dari berbagai pihak demi kesempurnaan tugas akhir ini.

Pekanbaru, 28 Juni 2011

Penulis

DAFTAR ISI

	Halaman
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN.....	iii
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL.....	iv
LEMBAR PERNYATAAN	v
LEMBAR PERSEMBAHAN.....	vi
ABSTRAK	vii
<i>ABSTRACT</i>	viii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR SIMBOL	xiii
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN.....	xvi
 BAB I PENDAHULUAN	
1.1 Latar Belakang	I-1
1.2 Rumusan Masalah	I-2
1.3 Batasan Masalah	I-2
1.4 Tujuan dan Manfaat	I-3
1.5 Sitematika Penulisan	I-3
 BAB II LANDASAN TEORI	
2.1 Pendapatan Domestik Regional Bruto	II-1
2.2 Analisis Multivariat.....	II-2
2.2.1 Analisis Komponen Utama	II-2
2.2.2 Analisis <i>Cluster</i>	II-8
2.2.2.1 Ukuran Kemiripan Objek.....	II-8
2.2.2.2 Metode Analisis <i>Cluster</i>	II-9
2.2.2.3 Metode Perbaikan Jarak.....	II-10

2.2.2.4 Dendogram.....	II-12
BAB III METODOLOGI PENELITIAN	
3.1 Jenis dan Sumber Data	III-1
3.2 Metode Analisis Data	III-2
3.2.1 Analisis Komponen Utama	III-2
3.2.2 Analisis <i>Cluster</i>	III-2
3.2.3 Dendogram.....	III-3
BAB IV PEMBAHASAN DAN HASIL	
4.1 Statistik Deskriptif	IV-1
4.2 Korelasi Antar Variabel.....	IV-2
4.3 Analisis Komponen Utama.....	IV-3
4.4 Analisis <i>Cluster</i>	IV-5
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan.....	V-1
5.2 Saran.....	V-2
DAFTAR PUSTAKA	
LAMPIRAN	
DAFTAR RIWAYAT HIDUP	

DAFTAR SIMBOL

X_i	: Variable asal
Y_i	: Variable baru
var	: Variansi
Σ	: Matrik kovarian
$L(x)$: Fungsi langrange
λ	: Nilai eigen
I	: Matrik identitas
δ	: Gamma
Z	: Z skor
σ	: Simpangan baku
μ	: Rata-rata
L	: Matriks diagonal
L^{-1}	: Invers matriks L
X^T	: Transpose matriks X
d	: Jarak

DAFTAR TABEL

Tabel		Halaman
4.1	Korelasi antar variabel yang diamati.....	IV-3
4.2	Skor komponen utama	IV-4
4.3	Proporsi keragaman dan keragaman kumulatif masing-masing komponen utama	IV-4
4.4	Proses pembentukan <i>cluster</i>	IV-6
4.5	Kabupaten/kota berdasarkan hasil analisis <i>cluster</i>	IV-8

DAFTAR LAMPIRAN

Lampiran	Halaman
A Data pendapatan domestik regional bruto atas dasar harga berlaku menurut lapangan usaha 11 kabupaten/kota di propinsi Riau Tahun 2010 (jutaan rupiah)	A-1
B Data pengamatan yang sudah dibakukan.....	B-1
C Proses analisis komponen utama dengan penguraian nilai singular	C-1
D Jarak <i>euclid</i>	D-1
E Tahap-tahap pembentukan <i>cluster</i>	E-1
F Nilai rata-rata variabel pada setiap <i>cluster</i>	F-1

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perencanaan pembangunan ekonomi suatu daerah memerlukan beragam data statistik untuk dasar penentuan strategi dan kebijaksanaan, agar sasaran pembangunan dapat dicapai dengan tepat. Strategi dan kebijaksanaan pembangunan ekonomi yang telah diambil pada masa-masa yang lalu perlu dimonitor dan dilihat hasil-hasilnya (BPS Riau, 2011).

Tingkat pertumbuhan pendapatan masyarakat suatu wilayah (region) dapat diketahui pada pendapatan regional secara berkala yang dapat digunakan sebagai bahan perencanaan pembangunan regional/daerah khususnya di bidang ekonomi. Angka-angka Pendapatan Domestik Regional Bruto (PDRB) merupakan suatu indikator berupa data agregat (kumpulan data) yang dapat dipakai untuk mengukur tingkat pertumbuhan ekonomi suatu daerah. Pemerintah daerah menaruh perhatian besar dalam pembangunan sistem data untuk memonitor perkembangan kemajuan disegala bidang (BPS Riau, 2011).

Dalam analisis statistik, sektor lapangan usaha dinyatakan sebagai variabel, semakin banyak jenis yang menjadi sektor, akan semakin rumit analisis yang harus dilakukan. Dalam penulisan ini, akan digunakan analisis multivariat yang merupakan analisis yang cocok untuk meringkas data dengan variabel yang banyak. Beberapa analisis dalam analisis multivariat yang dapat digunakan untuk memahami dan mempermudah interpretasi data multivariat diantaranya adalah analisis *cluster* (Purwaningsih, 2007).

Analisis *cluster* bertujuan menempatkan sekumpulan objek kedalam dua atau lebih grup berdasarkan kesamaan objek atas dasar berbagai karekteristik. Banyak objek yang dapat dikelompokkan dengan analisis *cluster* diantaranya adalah produk (barang dan jasa), benda, manusia (responden, konsumen) (Supranto, 2004).

Penelitian Hasma Chairani (2004) yaitu mengelompokkan Kabupaten/Kota di Sumatera Barat berdasarkan PDRB Tahun 2002 serta menentukan karakteristik dari masing-masing kelompok. Hasil penelitiannya menghasilkan 5 gerombol, yaitu: pada gerombol I terdiri dari 5 kabupaten/kota yaitu Kabupaten Mentawai, Kota Solok, Kota Padang Panjang, Kota Bukittinggi dan Kota Payakumbuh. Pada gerombol II terdiri dari 7 kabupaten/kota yaitu Kabupaten Pesisir Selatan, Kabupaten Solok, Kabupaten Tanah Datar, Kabupaten Padang Pariaman, Kabupaten Agam, Kabupaten 50 Kota dan Kabupaten Pasaman, sedangkan pada gerombol III, IV dan V masing-masingnya terdiri dari satu kabupaten/kota yaitu Kabupaten Sawahlonto/Sijunjung, Kota Padang dan Kota Sawahlonto.

Provinsi Riau memiliki sumber daya alam, baik kekayaan yang terkandung di perut bumi, berupa minyak bumi dan gas, maupun kekayaan hasil hutan dan perkebunannya, belum lagi kekayaan sungai dan lautnya. Seiring otonomi daerah, secara bertahap mulai diterapkan sistem bagi hasil atau perimbangan keuangan. Aturan baru dari pemerintahan reformasi, memberi batasan dan aturan tegas mengenai kewajiban penanam modal, pemanfaatan sumber daya dan bagi hasil dengan lingkungan sekitar (BPS Riau, 2011).

Berdasarkan uraian diatas, maka penulis tertarik untuk mengelompokkan setiap Kabupaten/Kota di Riau berdasarkan PDRB dengan menggunakan analisis *cluster*. Adapun judul tugas akhir yang diajukan adalah: **“Analisis *Cluster* Kabupaten/Kota di Riau berdasarkan Pendapatan Domestik Regional Bruto (PDRB) Tahun 2010”**.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang tersebut dapatlah dirumuskan masalah dalam penelitian ini yaitu: “Bagaimana pengklasifikasian atau pengelompokkan setiap Kabupaten/Kota di Riau berdasarkan PDRB Tahun 2010”.

1.3 Batasan Masalah

Dalam tugas akhir ini penulis akan menggunakan beberapa metode yaitu analisis komponen utama untuk mengubah variabel asal menjadi variabel baru,

kemudian dilanjutkan dengan analisis *cluster* untuk mengklasifikasi setiap Kabupaten/Kota di Riau berdasarkan PDRB dengan membatasi pencarian jarak dengan menggunakan metode *eucledian* dan perbaikan jarak dengan metode pautan tunggal/*single linkage*.

Data dalam penelitian ini diambil dari BPS (Badan Pusat Statistik) Riau, yaitu data tentang pendapatan domestik regional bruto atas dasar harga berlaku menurut lapangan usaha di Riau pada Tahun 2010.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan Penelitian

Tujuan yang hendak dicapai dalam penelitian ini yaitu:

1. Mengelompokkan Kabupaten/Kota di Riau berdasarkan PDRB menurut sektor lapangan usaha.
2. Mengetahui karakteristik dari setiap kelompok yang dihasilkan.

1.4.2 Manfaat Penelitian

Berdasarkan rumusan masalah dan tujuan penelitian, manfaat yang dapat diperoleh adalah :

1. Penulis dapat mengetahui ciri PDRB dari setiap Kabupaten/Kota di Riau.
2. Dapat menambah wawasan penulis mengenai analisis multivariat.
3. Dapat dijadikan sebagai bahan masukan bagi pemerintah mengenai kondisi perekonomian, sebagai bahan evaluasi dan rencana pembangunan selanjutnya.

1.5 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini mencakup lima bab, yaitu:

BAB I Pendahuluan

Bab ini berisi tentang latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II Landasan Teori

Bab ini berisi tentang informasi teori-teori yang digunakan dalam penulisan ataupun metode yang dipakai. Teori yang akan digunakan dalam penulisan ini yaitu analisis multivariat, analisis komponen utama, analisis *cluster* berhirarki penggabungan seperti jarak *euclidean* sebagai ukuran ketakmiripannya dan metode pautan tunggal sebagai metode perbaikan jarak.

BAB III Metodologi Penulisan

Bab ini berisikan langkah-langkah dalam menyelesaikan analisis komponen utama dan analisis *cluster*.

BAB IV Hasil dan Pembahasan

Bab ini berisikan tentang hasil dan pembahasan dari penulisan ini

BAB V Penutup

Bab ini berisikan tentang kesimpulan dan saran.

BAB II

LANDASAN TEORI

2.1 Pendapatan Domestik Regional Bruto

Perencanaan pembangunan ekonomi suatu daerah memerlukan beragam data statistik untuk dasar penentuan strategi dan kebijaksanaan, agar sasaran pembangunan dapat dicapai dengan tepat. Strategi dan kebijaksanaan pembangunan ekonomi yang telah diambil pada masa-masa yang lalu perlu dimonitor dan dilihat hasil-hasilnya. Berbagai data statistik yang merupakan ukuran kuantitas mutlak diperlukan untuk memberikan gambaran tentang keadaan pada masa yang lalu dan masa kini, serta sasaran-sasaran yang akan dicapai pada masa yang akan datang (BPS Sekadau, 2008).

Perhitungan dilakukan dengan dua cara:

1. PDRB atas dasar harga berlaku

Pada PDRB atas dasar harga berlaku, metode penghitungan dilakukan dengan menggunakan data yang bersumber dari daerah yang dapat memperlihatkan karakteristik sosial ekonomi setiap daerah. Jumlah nilai produk barang dan jasa akhir, dihasilkan oleh berbagai unit produksi didalam suatu region dalam jangka waktu tertentu (satu tahun). Unit-unit tersebut di atas dalam penyajiannya dikelompokkan menjadi 9 lapangan usaha yaitu (Badan Perencanaan Pembangunan Kabupaten Malang, 2007):

1. Pertanian
2. Pertambangan dan Penggalian
3. Industri pengolahan
4. Listrik, gas dan air bersih
5. Bangunan
6. Perdagangan, hotel, dan restoran
7. Pengangkutan dan komunikasi
8. Keuangan, persewaan, dan jasa perusahaan, dan
9. Jasa-jasa termasuk jasa pelayanan pemerintah.

2. PDRB atas dasar harga konstan

PDRB atas dasar harga konstan sangat penting untuk melihat perkembangan riil dari tahun ke tahun dari setiap agregat (kumpulan data) ekonomi yang diamati. Agregat yang dimaksud dapat merupakan produk Domestik Regional Bruto secara keseluruhan maupun nilai tambah sektoral. Pada dasarnya dikenal empat cara penghitungan nilai tambah sektoral atas dasar harga konstan, yaitu: revaluasi, ekstrapolasi, deflasi dan deflasi berganda (Badan Perencanaan Pembangunan Kabupaten Malang, 2007).

2.2 Analisis Multivariat

Analisis multivariat merupakan salah satu jenis analisis statistik yang digunakan untuk menganalisis data yang terdiri dari banyak variabel, baik variabel bebas (*independent variables*) maupun banyak variabel tak bebas (*dependent variabel*). Data multivariat adalah data yang dikumpulkan dari dua atau lebih observasi dengan mengukur observasi tersebut dengan beberapa karakteristik (Wijaya, 2010).

2.2.1 Analisis Komponen Utama (*Principal Component Analysis*)

Analisis Komponen Utama (AKU) secara teknis merupakan suatu teknik mereduksi data multivariat yang mengubah (mentransformasikan) variabel asal (suatu matriks data/asli) menjadi variabel baru yang merupakan kombinasi linier dari variabel asal (suatu set kombinasi linier yang lebih sedikit), akan tetapi menyerap sebagian besar jumlah varian dari data awal. Tujuan utamanya menjelaskan sebanyak mungkin jumlah varian data asli dengan sedikit mungkin komponen utama (Supranto, 2004).

Variabel baru ini memanfaatkan informasi dari variabel asal. Variabel baru ini harus bersifat saling orthogonal, jika dua variabel saling orthogonal maka variabel tersebut tidak berkorelasi (korelasi 0), berarti kuat hubungan yang terjadi antara variabel-variabel sama dengan nol. Ragam suatu variabel merupakan sifat yang penting yang digunakan dalam analisis, makin beragam suatu variabel, makin besar perannya dalam pemisahan antar objek (Repository.usu.ac.id, 2011).

Andaikan $X = [X_1, X_2, X_3, \dots, X_p]$ merupakan vektor variabel asal yang diamati dengan matrik kovarian $\Sigma(\sigma_{ij})$, maka komponen utama pertama yang dilambangkan dengan Y_1 didefinisikan sebagai :

$$Y_1 = \sum a_{ij} X_j = a_1^T X \quad (2.1)$$

dimana:

$$\begin{aligned} Y_1 &= a_{11}X_1 + a_{12}X_2 + a_{13}X_3 + \dots + a_{1p}X_p \\ &\vdots \\ Y_p &= a_{p1}X_1 + a_{p2}X_2 + a_{p3}X_3 \dots + a_{pp}X_p = a_p^T X \end{aligned} \quad (2.2)$$

dengan:

Y_p = variabel acak hasil transformasi dari variabel asal ke variabel baru

X_i = variabel acak asal

X = matriks variabel acak asal

dalam bentuk matrik :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \dots & a_{pp} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} \quad (2.3)$$

yang memaksimumkan ragam Y_1 , yaitu $a_1' \Sigma a_1$, dengan kendala $a_1^T a_1 = 1$. Kalau matrik kovarian dari vektor x adalah Σ , maka variansi Y_i dirumuskan sebagai:

$$var(Y_i) = a_i^T \Sigma a_i \quad (2.4)$$

Permasalahan transformasi adalah bagaimana memilih koefisien dari kombinasi linier tersebut sehingga :

$$var(Y_1) > var(Y_2) > \dots > var(Y_p) \quad (2.5)$$

Dengan menggunakan fungsi Langrange diperoleh a_1, a_2, \dots, a_p sebagai eigen vektor yang berpadanan dengan eigen value $\lambda_1, \lambda_2, \dots, \lambda_p$ dari matrik kovarian Σ . Bentuk umum persamaan Langrange adalah :

$$L(x) = f(x) - \lambda[g(x) - c] \quad (2.6)$$

dengan :

$f(x)$ = fungsi tujuan

$g(x)$ = fungsi kendala

c = konstanta

Penentuan a_1

Masalah optimum :

Maks $var(Y_1) = a_1^T \Sigma a_1$

Kendala $a_1^T a_1 = 1$ atau $a_1^T a_1 - 1 = 0$

Melalui Langrange, fungsi yang dimaksimumkan adalah:

$$L(a_1) = a_1^T \Sigma a_1 - \lambda_1(a_1^T a_1 - 1)$$

Pengoptimuman dilakukan dengan cara menurunkan fungsi L terhadap variabel-variabel yang dicari, dan diperoleh:

$$\begin{aligned}\frac{\partial L}{\partial a_1} &= \frac{\partial}{\partial a_1}(a_1^T \Sigma a_1 - \lambda_1(a_1^T a_1) + \lambda_1) \\ &= \frac{\partial}{\partial a_1}(a_1^T \Sigma a_1) - \frac{\partial}{\partial a_1}(\lambda_1(a_1^T a_1)) + \frac{\partial}{\partial a_1}(\lambda_1) \\ &= 2 \Sigma a_1 - 2 \lambda_1 a_1 + 0 \\ &= 2(\Sigma - \lambda_1 I) a_1 \\ &= (\Sigma - \lambda_1 I) a_1\end{aligned}$$

Supaya $\frac{\partial L}{\partial a_1} = 0$, maka harus $(\Sigma - \lambda_1 I) a_1 = 0$

sehingga:

$$(\Sigma - \lambda_1 I) a_1 = 0 \tag{2.7}$$

dengan :

Σ = matriks kovarian

λ_1 = akar karakteristik dari Σ

a_1 = vektor karakteristik padanan akar karakteristik λ_1

I = matriks identitas.

Persamaan 2.7 diatas dapat disederhanakan menjadi:

$$\begin{aligned}(\Sigma - \lambda_1 I) a_1 &= 0 \\ \Sigma a_1 - \lambda_1 I a_1 &= 0 \\ \Sigma a_1 &= \lambda_1 a_1\end{aligned} \tag{2.8}$$

Ini berarti a_1 adalah vektor eigen dari matriks Σ dengan nilai eigen λ . Jika persamaan 2.8 dikalikan dengan a_1^T , maka:

$$\begin{aligned} a_1^T \Sigma a_1 &= a_1^T \lambda_1 a_1 \\ &= \lambda_1 a_1^T a_1 \end{aligned}$$

dengan kendala $a_1^T a_1 = 1$, maka $\text{var}(Y_1) = a_1^T \Sigma a_1 = \lambda_1$, sehingga $\text{var}(Y_1)$ menjadi maksimum jika λ_1 adalah maksimum. Komponen utama kedua dilambangkan dengan Y_2 , didefinisikan sebagai berikut:

$$Y_2 = a_2^T X \quad (2.9)$$

yang memaksimumkan ragam Y_2 , dengan kendala $a_2^T a_2 = 1$ dan $a_1^T a_2 = 0$.

Penentuan a_2

Masalah optimum:

$$\begin{aligned} \text{Maks} &: \text{var}(Y_2) = a_2^T \Sigma a_2 \\ \text{Kendala} &: a_2^T a_2 = 1 \text{ atau } a_2^T a_2 - 1 = 0 \\ & a_1^T a_2 = 0 \end{aligned}$$

dengan menggunakan fungsi Langrange, maka fungsi yang dimaksimumkan adalah:

$$L(a_2) = a_2^T \Sigma a_2 - \lambda_2(a_2^T a_2 - 1) - \delta(a_1^T a_2)$$

dengan menurunkan fungsi L terhadap variabel yang akan ditentukan, maka pengoptimuman dapat dilakukan, yaitu:

$$\begin{aligned} \frac{\partial L}{\partial a_2} &= \frac{\partial}{\partial a_2} (a_2^T \Sigma a_2 - \lambda_2(a_2^T a_2 - 1) - \delta(a_1^T a_2)) \\ 0 &= \frac{\partial}{\partial a_2} (a_2^T \Sigma a_2) - \frac{\partial}{\partial a_2} (\lambda_2(a_2^T a_2)) + \frac{\partial}{\partial a_2} (\lambda_2) - \frac{\partial}{\partial a_2} (\delta a_1^T a_2) \\ 0 &= 2 \Sigma a_2 - 2 \lambda_2 a_2 + 0 - \delta a_1 \end{aligned} \quad (2.10)$$

Kemudian dikalikan dengan a_1^T , maka:

$$\begin{aligned} 2a_1^T \Sigma a_2 - 2 \lambda_2 a_1^T a_2 + 0 - \delta a_1 a_1^T &= 0 \\ 2a_1^T \Sigma a_2 - 0 + 0 - \delta &= 0 \\ 2a_1^T \Sigma a_2 &= \delta \end{aligned}$$

Dari persamaan 2.8, dikalikan dengan a_2^T , maka diperoleh:

$$\begin{aligned} \Sigma a_1 &= \lambda_1 a_1 \\ \Sigma a_1 a_2^T &= \lambda_1 a_1 a_2^T \\ &= 0 \end{aligned}$$

Oleh karena $\sum a_1 a_2^T = 0$, maka $\delta = 0$.

Persamaan 2.10 menjadi

$$2 \sum a_2 - 2 \lambda_2 a_2 + 0 - \delta a_1 = 0$$

$$2 \sum a_2 - 2 \lambda_2 a_2 = 0$$

$$(\sum - \lambda_2) a_2 = 0$$

$$(\sum - \lambda_2 I) a_2 = 0$$

Dengan mengalikan a_2^T , maka:

$$a_2^T (\sum - \lambda_2) a_2 = 0$$

$$a_2^T \sum a_2 - \lambda_2 a_2^T a_2 = 0$$

$$a_2^T \sum a_2 = \lambda_2$$

Ini berarti bahwa vektor a_2 merupakan vektor eigen dari \sum yang bersesuaian dengan nilai eigen terbesar kedua, yaitu λ_2 .

Peranan komponen utama dapat diukur dengan besarnya persentase keragaman total yang mampu diterangkan oleh komponen utama tersebut yaitu :

$$P_{kj} = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad (2.11)$$

dengan :

P_{kj} = Peranan komponen utama ke- j

λ_j = Akar karakteristik ke- j

Matriks kovarian \sum yang digunakan dalam masalah ini jika variabel yang diamati ukurannya pada skala dengan perbedaan yang sangat lebar atau jika satuan ukurannya tidak sama . Bila variabel yang diamati ukurannya pada skala dengan perbedaan yang sangat lebar atau jika satuan ukurannya tidak sama, maka variabel tersebut perlu dibakukan (*standardized*) sehingga komponen utama ditentukan dari variabel baku (Z) (Statistikaterapan.files.wordpress.com, 2008). Variabel baku (Z) dapat ditentukan dengan menggunakan persamaan sebagai berikut (Irianto, 2009):

$$Z = \frac{X - \mu}{\sigma} \quad (2.12)$$

dengan :

X = nilai variabel

σ = simpangan baku

μ = nilai rata-rata

Matriks skor komponen utama dapat diperoleh dari penguraian nilai singular, dari data yang telah dibakukan dibentuk matriks X , dapat diuraikan menjadi:

$$X_{n \times p} = U_{n \times r} L_{r \times r} A_{r \times p}^T$$

Dalam hal ini

$A = [a_1 \ a_2 \ \dots \ a_k]$; $AA^T = 1$ $A^T A = 1$, dengan a adalah vektor karakteristik dari matriks $X^T X$. L = matriks diagonal yang unsur diagonal utamanya adalah akar dari akar karakteristik tak nol dari matriks $X^T X$. Dalam hal ini λ_1 adalah akar karakteristik yang bersesuaian dengan vektor karakteristik a_1 . Berikut adalah matriks diagonal L :

$$L = \begin{bmatrix} \sqrt{\lambda_1} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & 0 & \dots & 0 \\ 0 & 0 & \sqrt{\lambda_3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sqrt{\lambda_r} \end{bmatrix}, \text{ dengan } k = 1, 2, \dots, p \text{ dan } \lambda_1 > \lambda_2 > \dots > \lambda_r$$

matriks U diperoleh dengan menggunakan rumus :

$$U = X A L^{-1}$$

nilai skor komponen utama diperoleh dengan menggunakan persamaan :

$$Y = XA \tag{2.13}$$

dengan :

Y = matriks skor komponen utama

A = matriks eigen vektor

X = matriks variabel asal

Selanjutnya, hasil dari analisis komponen utama dapat digunakan untuk perhitungan analisis selanjutnya, yaitu dalam analisis *cluster*.

2.2.2 Analisis Cluster

Analisis *cluster* merupakan teknik multivariat yang mempunyai tujuan utama mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Analisis *cluster* mengklasifikasikan objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam *cluster* yang sama (Wijaya, 2010).

Cluster-cluster yang terbentuk memiliki homogenitas (kesamaan) yang tinggi antar anggota dalam satu *cluster* dan heterogenitas (perbedaan) yang tinggi antar *cluster* yang satu dengan *cluster* yang lain. Solusi analisis *cluster* secara keseluruhan bergantung pada variabel-variabel yang digunakan sebagai dasar untuk menilai kesamaan (Wijaya, 2010).

2.2.2.1 Ukuran Kemiripan Objek

Sesuai prinsip dasar *cluster* yaitu mengelompokkan objek yang mempunyai kemiripan, maka proses pertama adalah mengukur seberapa jauh ada kesamaan antar objek. Dengan memiliki sebuah ukuran kuantitatif untuk mengatakan bahwa dua objek tertentu lebih mirip dibandingkan dengan objek lain, akan mempermudah proses pengelompokan. Pengelompokan dilakukan berdasarkan kemiripan antar objek. Kemiripan diperoleh dengan meminimalkan jarak antar objek dalam kelompok dan memaksimalkan jarak antar kelompok (Purwaningsih, 2007).

Ada beberapa jarak yang biasa digunakan dalam analisis *cluster*, yaitu:

1. Jarak *Euclidean*

Jarak *euclidean* dapat digunakan jika variabel-variabel yang digunakan tidak terdapat korelasi dan memiliki satuan yang sama. Jarak *euclidean* diperoleh dengan rumus sebagai berikut (Simamora, 2005):

$$d_{ij} = \sqrt{\sum_{k=1}^n (v_{ik} - v_{jk})^2} \quad (2.14)$$

dengan :

d_{ij} = jarak *euclidean*

v_{ik}, v_{jk} = skor komponen utama ke- i dan ke- j pada variabel k ($k = 1, 2, \dots, n$)

2. Jarak Mahalanobis

Mahalanobis adalah suatu metode statistika yang digunakan untuk mendapatkan suatu data dengan jarak tertentu terhadap *mean* data tersebut. Metode ini memiliki karakteristik yaitu dilihat dari rumusnya metode ini memiliki nilai penguat yang fleksibel sehingga mudah disesuaikan dengan perubahan kondisi, hal ini yang menyebabkan metode ini lebih akurat dibanding metode lain, Penguat dari mahalanobis ini ada pada nilai kovariannya Berikut adalah persamaan dari mahalanobis (Aribi, 2007):

$$d_{ij}^2 = (X - \mu)^T \Sigma^{-1} (X - \mu) \quad (2.15)$$

dengan:

d_{ij}^2 = jarak mahalanobis

Σ^{-1} = matrik kovarian

$(X - \mu)^T$ = vektor transpose

dengan Σ adalah kovarian variabel ke- j dan ke- j' sehingga dapat dihitung dengan menggunakan rumus :

$$\Sigma = \frac{\sum (x_{ij} - \bar{x}_j)(x_{ij'} - \bar{x}_{j'})}{n-1} \quad (2.16)$$

2.2.2.2 Metode Analisis Cluster

Secara umum ada dua metode dalam analisis *cluster* yaitu :

1. Metode Hirarki

Metode hirarki dapat digunakan jika jumlah kelompok yang diinginkan tidak diketahui, dan metode ini biasanya digunakan dalam objek dengan pengamatan yang tidak besar. Metode hirarki dibagi menjadi 2 (Simamora, 2005):

a. Penggabungan (*agglomerative*)

Dimulai dengan menempatkan objek ke dalam *cluster-cluster* yang berbeda, lalu mengelompokkan objek secara bertahap ke dalam *cluster-cluster* yang lebih besar. Artinya, pertama, setiap objek diperlakukan sebagai *cluster*. Jadi kalau ada 20 objek, pada tahap pertama, ada 20 *cluster*. Lalu, secara bertahap dilakukan pengelompokan dengan memeriksa satu pasangan objek paling mirip, lalu keduanya menjadi satu *cluster*. Jadi, pada tahap kedua, jumlah *cluster*

menjadi 19. Proses ini akan berjalan terus sehingga pada tahap terakhir, yaitu pada tahap ke-19, semua objek menjadi satu *cluster*.

b. Pemecahan (*divisive*)

Proses pemecahan merupakan kebalikan dari metode penggabungan. Metode ini dimulai dengan menempatkan semua objek sebagai satu *cluster*. Selanjutnya, secara bertahap, objek-objek dipisahkan ke dalam *cluster-cluster* yang berbeda, dua *cluster*, tiga *cluster*, dan seterusnya, sampai semua objek menjadi *cluster* sendiri-sendiri. Jadi kalau ada 20 objek, pada tahap terakhir akan ada 20 *cluster*.

2. **Metode Tak Hirarki**

Metode ini digunakan jika banyaknya kelompok sudah diketahui dan biasanya metode ini dipakai untuk mengelompokkan data yang berukuran besar (Statistikaterapan.files.wordpress.com, 2008).

Metode ini meliputi (Wahana_statistika.com, 2009):

a. *Sequential threshold*

Metode ini dimulai dengan memilih bakal *cluster* dan menyertakan seluruh objek dalam jarak tertentu. Jika seluruh objek dalam jarak tersebut disertakan, bakal *cluster* kedua terpilih, kemudian proses terus berlangsung seperti sebelumnya.

b. *Parallel threshold*

Metode ini memilih beberapa bakal *cluster* secara simultan pada permulaannya dan menandai objek-objek dengan jarak permulaan ke bakal terdekat.

c. *Optimizing partitioning*

Metode ketiga ini mirip dengan kedua metode sebelumnya kecuali pada penandaan ulang terhadap objek-objek.

2.2.2.3 Metode Perbaikan Jarak

Metode perbaikan jarak yang dapat digunakan pada metode hirarki adalah (Supranto, 2004):

1. Metode pautan tunggal (*single linkage*)

Cluster dibentuk dari individu objek dengan jelas menggabungkan jarak terdekat. Pada setiap tahap, setelah terbentuk *cluster* baru (*UV*), maka jarak antara (*UV*) dan *cluster* lainnya, misal *W* adalah :

$$d_{(UV)W} = \min \{d_{UW}, d_{VW}\} \quad (2.17)$$

dengan :

$d_{(U,W)}$ = jarak antara *cluster U* dan *W*

$d_{(V,W)}$ = jarak antara *cluster V* dan *W*

$d_{(UV,W)}$ = jarak antara *cluster (UV)* dan *W*

2. Metode pautan lengkap (*complete linkage*)

Metode pautan lengkap dilakukan sama seperti metode pautan tunggal, dengan suatu pengecualian yaitu, pada setiap tahap jarak antara *cluster* ditentukan oleh jarak antara dua objek, satu dari setiap *cluster* yang paling jauh. Setelah *cluster U* dan *V* digabung menjadi *cluster (UV)* jarak antara *cluster (UV)* dan *cluster* lain, misal *W* adalah :

$$d_{(UV)W} = \max \{d_{UW}, d_{VW}\} \quad (2.18)$$

dengan :

$d_{(U,W)}$ = jarak antara *cluster U* dan *W*

$d_{(V,W)}$ = jarak antara *cluster V* dan *W*

$d_{(UV,W)}$ = jarak antara *cluster (UV)* dan *W*

3. Metode pautan rata-rata (*average linkage*)

Dalam pautan rata-rata menggunakan jarak antara dua *cluster* sebagai rata-rata jarak antara semua pasangan objek dimana satu anggota pasangan menjadi milik setiap *cluster*. Terlebih dahulu tentukan matrik jarak $D = [d_{ik}]$, yang digunakan untuk menemukan objek yang berdekatan, misal *U* dan *V*. Selanjutnya,

objek tersebut digabungkan sehingga membentuk *cluster* (UV), dengan ukuran jarak antara *cluster* (UV) dan *cluster* W adalah :

$$d_{(UV)W} = \frac{d_{(U,W)} + d_{(V,W)}}{N_{UV}} \quad (2.19)$$

dengan :

$N_{(UV)}$ = jumlah pengamatan pada *cluster* (UV)

$d_{(U,W)}$ = jarak antara *cluster* U dan W

$d_{(V,W)}$ = jarak antara *cluster* V dan W

$d_{(UV,W)}$ = jarak antara *cluster* (UV) dan W .

2.2.2.4 Dendogram

Dendrogram merupakan diagram bercabang yang menggambarkan hirarki kategori berdasarkan derajat kesamaan sejumlah karakteristik. Hasil dari *cluster* dapat digambarkan dalam bentuk dendogram.

BAB III

METODOLOGI PENELITIAN

Penulisan skripsi ini menggunakan metode *research library* (penelitian kepustakaan) yang berguna untuk mengumpulkan data dan informasi yang diperlukan dalam penelitian. Data dan informasi berasal dari buku-buku bacaan yang ada kaitannya dengan pembahasan. Selain itu penulis juga menggunakan studi literatur yang dimulai dari mencari nilai korelasi antar variabel-variabel dari data yang diamati, kemudian dilakukan transformasi dengan analisis komponen utama, selanjutnya mengelompokkannya dengan menggunakan analisis *cluster*. Langkah-langkah yang akan digunakan adalah sebagai berikut :

3.1 Jenis dan Sumber Data

Penelitian ini menggunakan data PDRB dari 11 Kabupaten/Kota di Riau pada Tahun 2010 berdasarkan harga berlaku menurut 9 lapangan usaha. Sumber data diperoleh dari BPS Provinsi Riau. Kabupaten/Kota di Riau meliputi Kabupaten Bengkalis, Kota Dumai, Kabupaten Indragiri Hilir, Kabupaten Indragiri Hulu, Kabupaten Kampar, Kabupaten Kuantan Singingi, Kota Pekanbaru, Kabupaten Pelalawan, Kabupaten Rokan Hilir, Kabupaten Rokan Hulu dan Kabupaten Siak.

Variabel yang akan diamati adalah PDRB berdasarkan 9 lapangan usaha di Riau, yaitu sektor pertanian, sektor pertambangan dan penggalian, sektor pengolahan, sektor listrik, gas dan air bersih, sektor bangunan, sektor perdagangan, sektor pengangkutan dan komunikasi, sektor keuangan, persewaan dan jasa perusahaan dan sektor jasa-jasa.

3.2 Metode Analisis Data

3.2.1 Analisis Komponen Utama

Langkah-langkah yang akan dilakukan dalam menentukan analisis komponen utama adalah sebagai berikut:

1. Tranformasi data yang siap dianalisis ke dalam bentuk baku Z .
2. Hitung korelasi antar variabel yang diamati.

Variabel-variabel baru yang akan digunakan untuk analisis *cluster* tidak saling berkorelasi. Untuk itu dilakukanlah AKU untuk memperoleh skor komponen utama.

3. Proses AKU dilakukan dengan penguraian nilai singular.
 - Data yang telah dibakukan dibentuk matrik X
 - Matrik X ditranspos menjadi X^T
 - Kalikan matrik X^T dengan X sehingga diperoleh matrik $X^T X$
 - Peroleh vektor eigen A dan matrik akar dari nilai eigen L dari matrik $X^T X$.
 - Inverskan matrik L sehingga diperoleh matrik L^{-1} .
 - Kemudian peroleh matrik U . Matrik U diperoleh dengan persamaan:
$$U = XAL^{-1} \text{ karena } X = ULA^T$$
 - Matrik skor komponen utama dapat diperoleh dengan $Y = XA$.
Skor komponen utama yang menjadi variabel-variabel baru yang akan digunakan untuk melakukan analisis gerombol, variabel tersebut tidak saling berkorelasi.
4. Nilai keragaman, serta proporsi keragaman kumulatif dari masing-masing skor komponen utama.

3.2.2 Analisis Cluster

Langkah-langkah yang akan dilakukan dalam analisis *cluster* adalah:

1. Menentukan ukuran ketakmiripan antara dua objek, ukuran ini dihitung menggunakan jarak *euclidean*. Proses pertama adalah mengukur kesamaan antara objek, sehingga ukuran ketakmiripan adalah fungsi jarak antar objek. Perhitungan jarak *euclidean* dapat digunakan persamaan 2.14.

2. Membuat *cluster*.

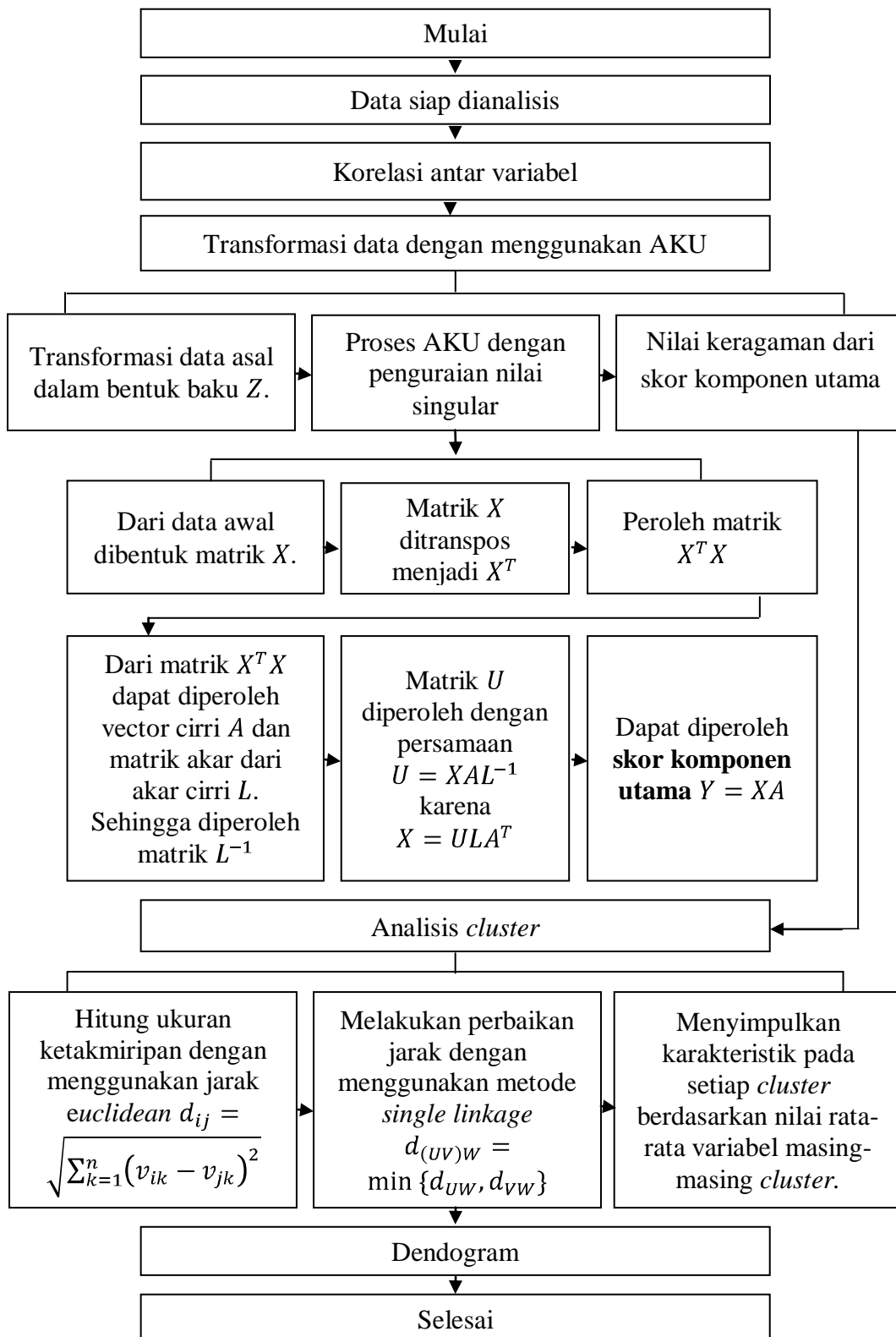
Metode yang digunakan untuk proses *cluster* secara hirarki adalah *single linkage*. Metode ini akan mengelompokkan dua objek yang mempunyai jarak terdekat. Jadi pada setiap tahapan, banyaknya *cluster* berkurang satu. Perhitungan perbaikan jarak *single linkage* dapat digunakan persamaan 2.17. Hasil dari *single linkage* dapat disajikan dalam bentuk dendogram atau diagram pohon.

3. Memberi nama spesifik setelah *cluster* terbentuk, untuk menggambarkan isi *cluster* tersebut. Dari dendogram dan *cluster-cluster* yang terbentuk dapat disimpulkan karakteristik dari setiap *cluster* yang terbentuk.

3.2.3 Dendogram

Hasil dari metode *cluster* dapat digambarkan dalam bentuk diagram pohon yang disebut dendogram. Jumlah *cluster* yang dihasilkan dibentuk berdasarkan pemotongan dendogram pada selisih jarak penggabungan terbesar.

Flow chart metodologi penelitian dengan AKU dan metode *cluster*, yaitu:



Gambar 3.1 Flow chart metodologi penelitian

BAB IV

PEMBAHASAN DAN HASIL

Bab IV ini akan dilakukan pembahasan analisis *cluster* untuk Kabupaten/Kota di Riau berdasarkan pendapatan domestik regional bruto (PDRB) Tahun 2010. Pertama, penjelasan mengenai statistik deskriptif kedua, menghitung nilai korelasi antar variabel ketiga, melakukan analisis komponen utama, selanjutnya melakukan analisis *cluster*.

4.1 Statistik Deskriptif

Data pendapatan domestik regional bruto atas dasar harga berlaku menurut lapangan usaha 11 kabupaten/kota di propinsi Riau Tahun 2010 (jutaan rupiah), dapat dilihat pada Lampiran A.

Pada lampiran A dapat dilihat bahwa pada PDRB sektor pertanian, jumlah yang tertinggi berada pada Kabupaten Indragiri Hilir, sedangkan yang terendah terdapat di Kota Dumai, dengan jumlah rata-rata pada PDRB sektor pertanian yaitu 6.710.773,15. Umumnya 11 Kabupaten/Kota yang ada di Riau PDRB sektor pertaniannya berada diatas rata-rata dan terdapat 4 Kabupaten yang PDRB sektor pertaniannya menjadi unggulan dibandingkan PDRB sektor lainnya.

Pada PDRB sektor pertambangan dan penggalian, yang menyumbang paling besar terdapat di Kabupaten Kampar, sedangkan penyumbang paling sedikit terdapat di Kota Dumai. Secara umum PDRB sektor pertambangan dan penggalian yang ada di 11 Kabupaten/Kota di Riau berada di bawah rata-rata. Terdapat 4 Kabupaten/Kota yang mengandalkan PDRB sektor pertambangan dan penggalian.

Selanjutnya, pada PDRB sektor industri dan pengolahan, Kabupaten Siak menempati posisi teratas dibandingkan Kabupaten/Kota lainnya, sedangkan posisi terendah terdapat di Kota Dumai. Secara umum PDRB sektor industri dan pengolahan dari 11 Kabupaten/Kota yang ada di Riau berada di bawah rata-rata. Terdapat 2 Kabupaten yang mengandalkan PDRB sektor industri dan pengolahan.

Pada PDRB sektor listrik, gas dan air bersih, hamper seluruh kabupaten/kota memerlukan perhatian khusus, karena PDRB sektor listrik, gas dan air bersih memiliki nilai paling rendah diantara sektor-sektor lainnya. Kota Pekanbaru memiliki nilai paling tinggi pada sektor listrik, gas dan air bersih. Nilai paling rendah berada di kabupaten Rokan Hulu dan rata-rata dari 11 kabupaten/kota, nilainya berada di bawah rata-rata.

PDRB sektor bangunan, PDRB sektor perdagangan, hotel dan restoran, PDRB sektor pengangkutan dan komunikasi, dan PDRB sektor keuangan, persewaan, dan jasa perusahaan. Pada keempat sektor ini kota Pekanbaru menempati posisi teratas dan Kabupaten Rokan Hulu berada pada posisi terendah. Sedangkan rata-rata dari 11 kabupaten/kota di Riau, hampir seluruh nilainya berada dibawah rata-rata. Kota Dumai mengandalkan PDRB sektor perdagangan, hotel dan restoran, karena nilainya lebih tinggi dibandingkan sektor lainnya.

Selanjutnya, PDRB sektor jasa-jasa, Kota Pekanbaru kembali menempati posisi teratas, sedangkan Kabupaten Pelalawan menempati posisi terendah. Rata-rata nilainya berada dibawah rata-rata.

4.2 Korelasi Antar Variabel

Langkah pertama yaitu menghitung nilai korelasi antar variabel yang akan diamati. Tabel 4.1 memuat suatu matriks korelasi, yang mempunyai dua bagian hubungan yang kuat. Bagian pertama meliputi hubungan antara variabel X_1 sampai dengan X_4 dengan koefisien korelasi -0,123; -0,187; -0,518, yang berada pada pojok kiri atas matriks korelasi. Bagian yang kedua berada pada bagian bawah sebelah kanan, meliputi hubungan antara variabel X_5 sampai dengan X_9 , dengan koefisien korelasi sebesar 0,967; 0,974; 0,967; 0,97; 0,956.

Untuk sementara, sebelum melakukan penelitian, dapat dianggap bahwa data asli sebanyak 9 variabel, cukup diwakili oleh dua faktor saja. Dari dua faktor atau komponen tersebut sudah menyerap sebagian besar varian yang terkandung dalam matriks korelasi sebagai data asli.

Analisis komponen utama menggunakan komponen dengan penyerapan varian data asli yang semakin mengecil (komponen pertama menyerap varian

terbesar pertama, komponen kedua menyerap varian terbesar kedua terhadap sisa varian dan seterusnya).

Tabel 4.1 Korelasi antar variabel yang diamati

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_2	-0,123							
X_3	-0,187	0,139						
X_4	-0,518	0,490	0,531					
X_5	-0,424	0,464	0,558	0,967				
X_6	-0,362	0,477	0,525	0,974	0,967			
X_7	-0,478	0,484	0,476	0,967	0,988	0,970		
X_8	-0,406	0,472	0,543	0,970	0,992	0,978	0,987	
X_9	-0,316	0,461	0,539	0,956	0,976	0,984	0,967	0,977

Keterangan:

X_1 = PDRB sektor pertanian

X_2 = PDRB sektor pertambangan dan penggalian

X_3 = PDRB sektor industri pengolahan

X_4 = PDRB sektor listrik, gas dan air bersih

X_5 = PDRB sektor bangunan

X_6 = PDRB sektor perdagangan, hotel, dan restoran

X_7 = PDRB sektor pengangkutan dan komunikasi

X_8 = PDRB sektor keuangan, persewaan, dan jasa perusahaan, dan

X_9 = PDRB sektor jasa-jasa termasuk jasa pelayanan pemerintah.

4.3 Analisis Komponen Utama

Skor komponen utama ditentukan dari variabel baru. Jika variabel yang diamati ukurannya pada skala dengan perbedaan yang sangat lebar atau satuan ukuran tidak sama, maka variabel tersebut perlu dibakukan, sehingga komponen utama ditentukan dari variabel baru. Data asal dapat dilihat pada Lampiran A, dan data yang telah dibakukan dapat dilihat pada Lampiran B.

Setelah data dibakukan, maka data tersebut akan digunakan sebagai variabel baru. Proses AKU dilakukan dengan penguraian nilai singular, proses

penguraian nilai singular dapat dilihat pada Lampiran C. Skor komponen utama inilah yang akan digunakan pada analisis *Cluster*.

Tabel 4.2 Skor komponen utama

Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9
-0.20128	0.24735	0.32267	-1.02378	0.50532	-0.03544	0.12519	-0.04315	0.01137
0.79377	0.13816	-0.08734	0.45884	-0.01070	0.03220	0.16825	0.11254	-0.07859
1.25086	-0.14730	0.14443	0.82157	-0.09595	0.08042	0.17101	-0.02841	0.05442
0.58047	-0.85318	-1.31587	0.83018	0.12993	-0.26204	-0.06614	-0.02465	-0.00527
0.43097	1.58662	1.59380	0.24744	-0.04599	-0.16430	-0.10302	0.05150	0.01456
1.01810	-0.41188	-0.17950	0.35456	0.26826	0.22496	-0.19605	0.07130	0.01647
0.99419	0.78594	-0.52320	-0.61885	-0.11938	0.06611	-0.08863	-0.09672	-0.06525
0.64656	0.64018	-1.47170	-1.00970	-0.24497	-0.01202	0.01769	0.06668	0.05129
0.62937	-2.10894	1.17620	-0.86311	-0.20113	-0.04316	-0.02598	0.01497	-0.00597
1.50523	0.15603	0.38771	0.59040	-0.12157	0.07606	0.00386	-0.11451	0.00713
-7.64825	-0.03299	-0.04722	0.21245	-0.06383	0.03722	-0.00619	-0.00955	-0.00018

Tabel 4.3 Proporsi keragaman dan keragaman kumulatif masing-masing komponen utama

Komponen Utama Ke-	Akar Ciri	Proporsi Keragaman	Keragaman Kumulatif
1	6.6358	0.737	0.737
2	0.9017	0.100	0.838
3	0.8412	0.093	0.931
4	0.5351	0.059	0.990
5	0.0489	0.005	0.996
6	0.0169	0.002	0.998
7	0.0136	0.002	0.999
8	0.0051	0.001	1.000
9	0.0017	0.000	1.000

Pada tabel 4.3, terlihat bahwa hanya komponen utama pertama yang memiliki nilai eigen lebih besar dari 1, yaitu 6.6358. Komponen pertama ini dapat menjelaskan 73,7% keragaman data. Komponen kedua memiliki nilai eigen 0.9017 dan dapat menjelaskan 10% keragaman, bersama dengan komponen pertama, keduanya mempresentasikan 83,8% dari keragaman total.

Komponen utama pertama dan komponen utama kedua yang mempresentasikan 83,8% keragaman total bisa dinilai telah cukup menangkap struktur data. Atau hanya dengan menggunakan komponen pertama telah cukup menangkap struktur data. Komponen-komponen lainnya memiliki proporsi keragaman yang kecil bisa dianggap tidak penting.

4.4 Analisis Cluster

Analisis *cluster* dilakukan dengan menggunakan metode berhirarki, karena jumlah *cluster* yang akan dibentuk tidak diketahui sebelumnya. Ukuran ketakmiripan yang digunakan adalah jarak Euclid. Metode perbaikan jarak yang digunakan dalam analisis ini adalah metode pautan tunggal. Proses yang dilakukan dalam analisis *cluster*, yaitu:

1. Menentukan Ukuran Ketakmiripan

Langkah awal dalam melakukan analisis *cluster* menghitung jarak masing-masing objek pengamatan, dapat digunakan Persamaan 2.11, sehingga diperoleh hasil seperti Lampiran D. Sebagai contoh jarak antara Kabupaten Bengkalis dengan Kabupaten Kuantan Sengingi sebesar 2,4659, sedangkan Kabupaten Bengkalis dengan Kabupaten Indragiri Hulu sebesar 1.9166. Hal ini menunjukkan bahwa Kabupaten Bengkalis lebih mirip karakteristiknya dengan Kabupaten Indragiri Hulu. Demikian seterusnya, semakin kecil nilai jarak antara dua kabupaten/kota, maka semakin mirip karakteristik kedua kabupaten/kota tersebut.

2. Pembentukan *Cluster*

Proses lengkap pembentukan *cluster* dengan analisis *cluster* dapat dilihat pada Lampiran E. Dengan tahapan sebagai berikut:

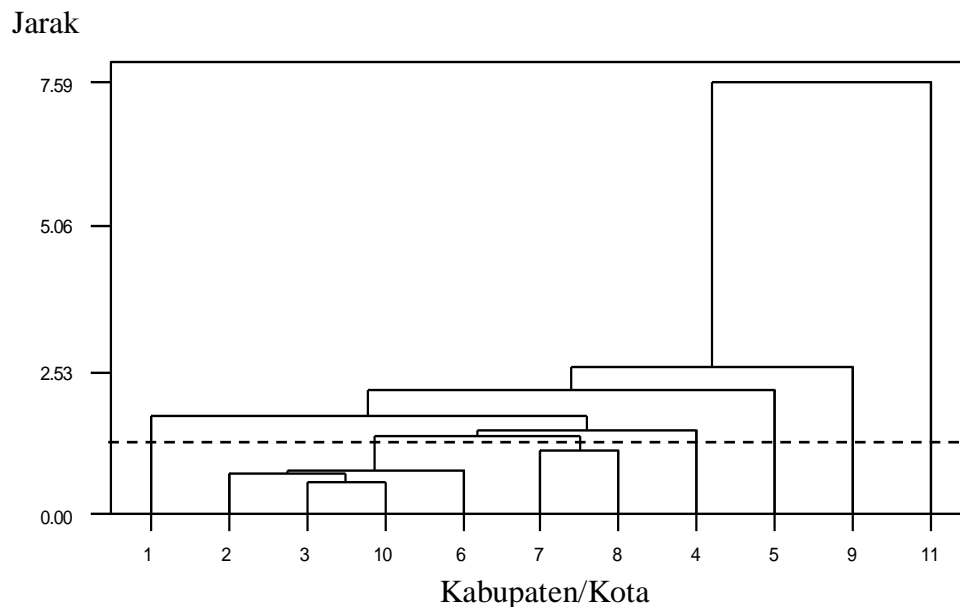
- Pada tahap 1, terbentuk satu *cluster* antara Kabupaten Kuantan Sengingi (3) dengan Kabupaten Rokan Hulu (10). Proses pengelompokan dimulai dengan jarak antar objek yang paling dekat, yaitu jarak antara Kabupaten Kuantan Sengingi dengan Kabupaten Rokan Hulu yang memiliki jarak terdekat dari 55 kombinasi jarak antar pengamatan yang ada. Selanjutnya lakukan perbaikan jarak dengan metode pautan tunggal, dapat digunakan Persamaan 2.17, sehingga diperoleh jarak Euclid yang baru. Jumlah *cluster* pada awalnya 11, namun pada tahap pertama jumlah *cluster* menjadi 10, karena penggabungan *cluster* 3 dan *cluster* 10, yang menyebabkan jumlah *cluster* berkurang.
- Pada tahap 2, dengan melihat jarak Euclid yang baru pada tahap 1, dapat dilihat terbentuknya *cluster* antara Kabupaten Indragiri Hulu (2) dengan

Kabupaten Kuantan Sengingi (3) sebesar 0,723 yang menunjukkan dasarnya jarak terdekat antara Kabupaten Indragiri Hulu dengan kedua kabupaten sebelumnya (Kabupaten Kuantan Sengingi dan Kabupaten Rokan Hulu). Dengan terbentuknya *cluster* tersebut, maka sekarang jumlah objek dalam *cluster* terdiri dari 3 objek, yaitu Kabupaten Indragiri Hulu, Kabupaten Kuantan Sengingi dan Kabupaten Rokan Hulu. Kemudian lakukan lagi perbaikan jarak, sehingga diperoleh kembali jarak Euclid yang baru untuk digunakan pada tahap selanjutnya. Jumlah *cluster* pada tahap kedua 9 buah, ini karena penggabungan *cluster* 2 dan *cluster* 3 dari tahap pertama dari jumlah *cluster* pertama.

- Pada tahapan selanjutnya, proses *cluster* dilakukan seperti tahapan sebelumnya hingga semua objek berada dalam satu *cluster*. Secara ringkas hasil analisis *cluster* disajikan pada Tabel 4.4 berikut ini:

Tabel 4.4 Proses pembentukan *cluster*

Tahap	Jumlah <i>Cluster</i>	Jarak	<i>Cluster</i> yang digabung		<i>Cluster</i> baru	Jumlah objek dalam <i>cluster</i>
1	10	0,554	3	10	3	2
2	9	0,723	2	3	2	3
3	8	0,794	2	6	2	4
4	7	1,126	7	8	7	2
5	6	1,390	2	7	2	6
6	5	1,478	2	4	2	7
7	4	1,747	1	2	1	8
8	3	2,208	1	5	1	9
9	2	2,583	1	9	1	10
10	1	7,586	1	11	1	11



Gambar 4.1 Dendogram hasil pengklasteran 11 Kabupaten/Kota berdasarkan 9 lapangan usaha

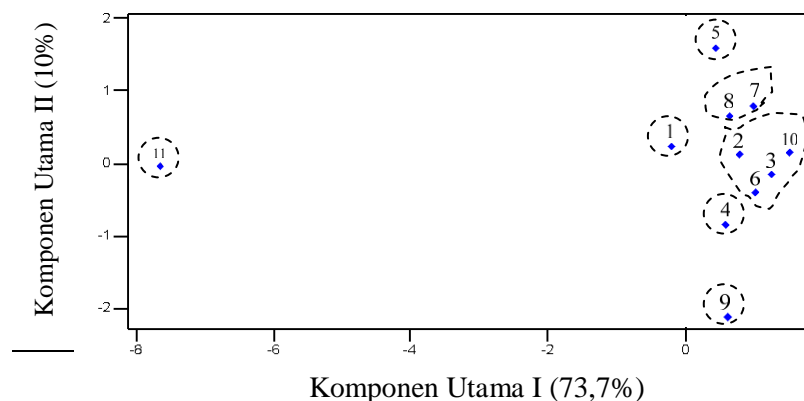
Dari dendogram dapat terlihat bahwa Kabupaten Kuantan Sengingi dan Kabupaten Rokan Hulu merupakan *cluster* berjarak sangat dekat. Ketika jarak yang diperbolehkan ditambah, maka Kabupaten Indragiri Hulu ditambahkan ke *cluster* Kabupaten Kuantan Sengingi dan Kabupaten Rokan Hulu. Perhatikan Kota Pekanbaru, pada *cluster* ini tidak menyatu sampai jarak diantara jarak terdekat telah meningkat secara tinggi. Pada akhirnya semua kabupaten/kota menyatu menjadi satu *cluster*.

Hasil pengelompokan diperoleh pemotongan pada selisih jarak 1,126 dengan 1,390. Pemotongan ini menghasilkan 7 *cluster* baru, secara rinci dapat dilihat pada Tabel 4.5 berikut ini:

Tabel 4.5 Kabupaten/kota berdasarkan hasil analisis *cluster*

<i>CLUSTER</i>	Kabupaten/Kota
I	Kabupaten Bengkalis
II	Kabupaten Indragiri Hulu, Kabupaten Kuantan Sengingi, Kabupaten Rokan Hulu, Kabupaten Rokan Hilir
III	Kabupaten Pelalawan, Kabupaten Siak
IV	Kabupaten Indragiri Hilir
V	Kota Dumai
VI	Kabupaten Kampar
VII	Kota Pekanbaru

Plot dua komponen utama berikut ini yaitu komponen utama satu dan dua memperlihatkan posisi masing-masing kabupaten/kota serta pengelompokannya, yang sudah mewakili 83,8% dari total keragaman data.



Gambar 4.2 Plot dua komponen utama pertama (Y1 dan Y2)

Keterangan gambar:

- | | |
|--------------------------|---------------------|
| 1. Kab. Bengkalis | 7. Kab. Pelalawan |
| 2. Kab. Indragiri Hulu | 8. Kab. Siak |
| 3. Kab. Kuantan Sengingi | 9. Kab. Kampar |
| 4. Kab. Indragiri Hilir | 10. Kab. Rokan Hulu |
| 5. Kota Dumai | 11. Kota Pekanbaru |
| 6. Kab. Rokan Hilir | |

3. Karakteristik *Cluster*

Setelah *cluster* terbentuk, tahap selanjutnya yaitu memberi gambaran mengenai isi *cluster* tersebut. Karakteristik pada tiap *cluster* dapat disimpulkan

berdasarkan nilai rata-rata variabel masing-masing *cluster*. Nilai rata-rata variabel pada setiap *cluster* dapat dilihat pada Lampiran F.

Dari ketujuh *cluster* yang terbentuk dapat diklasifikasikan sebagai berikut:

- *Cluster I*

Kabupaten Bengkalis yang berada pada *cluster I* memiliki rata-rata dibawah rata-rata umum, terlihat bahwa PDRB sektor pertambangan dan penggalian (X_2) memiliki nilai paling tinggi, sedangkan PDRB sektor listrik, gas dan air bersih (X_4) memiliki nilai paling rendah. Terdapat 3 sektor yang memiliki hasil diatas rata-rata umum.

- *Cluster II*

Empat kabupaten/kota yang berada pada *cluster II* memiliki rata-rata tertinggi diantara *cluster-cluster* lainnya. Pada *cluster II* nilai PDRB sektor pertanian (X_1) yang tertinggi, sedangkan sektor listrik, gas dan air bersih (X_4) masih berada ditingkat bawah. PDRB sebagian besar sektor-sektor pada *cluster II* berada dibawah rata-rata umum, hanya sektor pertanian (X_1) yang berada diatas rata-rata umum.

- *Cluster III*

Kabupaten Pelalawan dan Kabupaten Siak yang terdapat pada *cluster III* merupakan daerah-daerah yang perekonomiannya masih dibawah rata-rata. Sektor yang tertinggi yaitu sektor industri pengolahan (X_3), PDRB sektor listrik, gas dan air bersih (X_4), memiliki PDRB paling rendah diantara sektor-sektor yang ada. Secara umum sektor-sektor di *cluster III* memiliki nilai dibawah rata-rata umum,

- *Cluster IV*

Kabupaten Indragiri Hilir merupakan daerah yang terdapat pada *cluster IV* yang mengandalkan sektor pertanian (X_1). Jika dibandingkan dengan *cluster* lain, PDRB sektor pertanian (X_1) yang paling besar. PDRB sektor listrik, gas dan air bersih (X_4) memiliki PDRB paling rendah diantara sektor-sektor yang lain. *Cluster IV* memiliki rata-rata paling rendah dibandingkan dengan *cluster-cluster* lainnya.

- *Cluster V*

Kota Dumai, pada *cluster V*, PDRB sektor perdagangan, hotel dan restoran (X_6) memiliki nilai paling tinggi diantara sektor lainnya, sedangkan sektor pertambangan dan penggalian (X_2) memiliki nilai paling rendah dibandingkan dengan sektor-sektor lainnya dan *cluster-cluster* lainnya.

- *Cluster VI*

Kabupaten Kampar memiliki PDRB sektor pertambangan dan penggalian (X_2) yang tinggi, jauh melebihi nilai rata-rata umum, sehingga lebih unggul dibanding *cluster-cluster* yang lain. Sedangkan sektor listrik, gas dan air bersih (X_4) memiliki nilai paling rendah dibandingkan sektor lainnya. Sektor-sektor pada *cluster VI* berada dibawah rata-rata umum, hanya sektor pertambangan dan penggalian (X_2) dan sektor pertanian (X_1) yang berada diatas rata-rata umum.

- *Cluster VII*

Kota Pekanbaru yang menempati *cluster VII*. Berbagai sektor yang PDRBnya menempati urutan teratas dibanding PDRB sektor yang sama pada *cluster* lain. Sektor pertambangan dan penggalian (X_2) menempati nilai paling tinggi. PDRB sektor listrik, gas dan air bersih (X_4) memiliki PDRB paling rendah diantara sektor-sektor yang lain.

Hasil analisis *cluster* dari kabupaten/kota di Riau berdasarkan PDRB Tahun 2010 menurut lapangan usaha ini dapat memberikan gambaran mengenai kondisi perekonomian kabupaten/kota di Riau.

BAB V

KESIMPULAN DAN SARAN

Bab V ini berisi tentang kesimpulan dan saran yang diambil berdasarkan uraian dan pembahasan pada bab sebelumnya.

5.1 Kesimpulan

Kesimpulan yang diperoleh dari hasil penelitian ini, adalah terdapat 7 *cluster* yang dapat dibentuk berdasarkan PDRB Tahun 2010 pada Kabupaten/Kota di Riau, yaitu:

1. *Cluster I* : Kabupaten Bengkalis. Pada *cluster* ini, PDRB sektor pertambangan paling tinggi, sedangkan yang paling rendah PDRB sektor listrik, gas dan air bersih.
2. *Cluster II* : Kabupaten Indragiri Hulu, Kabupaten Kuantan Singingi, Kabupaten Rokan Hulu, Kabupaten Rokan Hilir. Pada *cluster* ini, PDRB sektor pertanian paling tinggi, sedangkan yang paling rendah PDRB sektor listrik, gas dan air bersih.
3. *Cluster III* : Kabupaten Pelalawan, Kabupaten Siak. Pada *cluster* ini, PDRB sektor industri paling tinggi, sedangkan yang paling rendah PDRB sektor listrik, gas dan air bersih.
4. *Cluster IV* : Kabupaten Indragiri Hilir. Pada *cluster* ini, PDRB sektor pertanian paling tinggi, sedangkan yang paling rendah PDRB sektor listrik, gas dan air bersih.
5. *Cluster V* : Kota Dumai. Pada *cluster* ini, PDRB sektor perdagangan, hotel dan restoran, sedangkan yang paling rendah PDRB sektor pertambangan dan penggalian.
6. *Cluster VI* : Kabupaten Kampar. Pada *cluster* ini, PDRB sektor pertambangan dan penggalian paling tinggi, sedangkan yang paling rendah PDRB sektor listrik, gas dan air bersih.

7. *Cluster VII*: Kota Pekanbaru. Pada *cluster* ini, PDRB sektor pertambangan dan penggalian paling tinggi, sedangkan yang paling rendah PDRB sektor pertanian.

5.2 Saran

Saran penulis bagi pihak yang membutuhkan informasi penelitian ini adalah:

1. Setelah mengetahui pengelompokan kabupaten/kota berdasarkan PDRBnya, diharapkan pemerintah dapat melihat gambaran karakteristik dari setiap kelompok yang dihasilkan sebagai bahan evaluasi dan perencanaan pembangunan ekonomi untuk masa yang akan datang.
2. Untuk peneliti selanjutnya diharapkan dapat menggunakan analisis statistik yang lain dalam pengolahan data.

DAFTAR PUSTAKA

- Anton, Howard. *Aljabar Linier Elementer*. Edisi ke-5. Erlangga, Jakarta. 1998.
- Aribi, Jauhar. *Mahalanobis Distance*. 2007 [Online] Available <http://jauhar-aribi.blogspot.com/2007/03/mahalanobis-distance.html>, diakses 04 Juni 2011.
- Badan Pusat Statistik. *Riau Dalam Angka 2010*. Badan Pusat Statistik, Pekanbaru. 2011.
- Badan Pusat Statistik. *Produk Domestik Regional Bruto Menurut Lapangan Usaha Kabupaten Sekadau 2004-2008*. 2008 [Online] Available http://sekadaukab.bps.go.id/pdf/pdrb_09_00.pdf, diakses 06 Maret 2011.
- Badan Perencanaan Pembangunan Kabupaten. *Produk Domestik Regional Bruto Kabupaten Malang Tahun 2007*. 2007 [Online] Available <http://bappekab.malangkab.go.id/admin/files/pendahuluan.pdf>, diakses 06 Maret 2011.
- Chairani, Hasma. “Pengerombolan Kabupaten/Kota di Sumatera Barat Berdasarkan PDRB”. *Tugas Akhir Mahasiswa Andalas*. Padang, 2004.
- Irianto, Agus. *Statistik Konsep Dasar dan Aplikasinya*. Kencana, Jakarta. 2009.
- Kuncoro, M. *Metode Riset Untuk Bisnis dan Ekonomi*. Erlangga, Jakarta. 2003.
- Repository.usu.ac.id. *Analisis Komponen Utama*. 2011 [Online] Available <http://repository.usu.ac.id/bitstream/123456789/16190/3/Chapter%20II.pdf>, diakses 09 Maret 2011.
- Sembiring, RK. *Analisis Regresi*. Edisi kedua. ITB. Bandung. 1995.
- Simamora, Bilson. *Analisis Multivariat Pemasaran*. PT Gramedia Pustaka Utama, Jakarta. 2005.
- Statistikaterapan.files.wordpress.com. *Analisis Kelompok (Cluster Analysis)*. 2008 [Online] Available <http://statistikaterapan.files.wordpress.com/2008/10/analisis-kelompok.doc>, diakses 11 Februari 2011.
- Statistikaterapan.files.wordpress.com. *Analisis Peubah Ganda (Multivariate Analysis)*. 2008 [Online] Available <http://statistikaterapan.files.wordpress.com/2008/10/analisis-akuklusterfaktor.doc>, diakses 11 Februari 2011.

Supranto, J. *Analisis Multivariat: Arti dan Interpretasi*. PT. Rineka Cipta, Jakarta. 2004

Wahana Statistika. *Analisis Cluster*. 2009 [Online] Available http://www.wahana-statistika.com/phocadownload/analisis_cluster.pdf, diakses 03 Februari 2011.

Wijaya, Tony. *Analisis Multivariat*. Universitas Atma Jaya Yogyakarta, Yogyakarta. 2010.